Actinobacteriophage Genome Annotation Submission Cover Sheet

This Cover Sheet will accompany each genome's annotation file(s) submission and succinctly describe the work that your students and you have done. This document ensures that the work done was as complete and thorough as it could be. Most important to the QC reviewer, denote where the trouble spots were in your annotation and how they were resolved.

Phage Name. **ObLaDi**
Your Name. Amanda Freise
Your Institution. UCLA
Your email. afreise@ucla.edu
Additional emails. (for correspondence). n/a

Describe any issues or specific genes that you would like to highlight for the QC reviewer. This includes any genes that you had questions about or received help with or that warrant further inspection in the QC review process. Include those genes that you deliberated on and/or want to strongly advocate for. If you contacted SMART, workshop facilitator, or a buddy school for help, please document.

Stop **26428**: possible Lysin A subunit functional call, but needs more/stronger evidence. Left as NKF for now. There are already two other Lysin A subunit calls which I feel confident about, and I haven't heard that 3 subunits is possible.
Stop **31259**: called as Cro, but would like review of functional evidence. HTH might be more appropriate if not enough to confirm Cro.
Stop **32287**: see above. This region is tricky. Cro/HTH/repressor. much overlapping evidence.
Stop **37515**: helicase loader. Has HHpred hits to support, but they also suggest initiator. Cluster member Aleemily also calls helicase loader.
Stop **49469**: called as DNA-binding protein in Cafasso, but not convinced by evidence here in ObLaDi. Cluster member Aleemily calls NKF.
Stop **47111** and **55293**: both called as DNAQ-like. Is it possible to have two? One may be in error if not. both have very compelling evidence
Stop **57064**: Good evidence to call as metallophosphoesterase, but it lacks the HEXXH motif that the approved functions list says is required. However, the example gene provided for this function (Luchador_60) also lacks this motif... Not sure how to proceed.
Stop **59914**: called as DNA-binding protein in Cafasso and Aleemily, but not convinced by evidence here in ObLaDi.
Stop **62369**: Some evidence for toxin gene, but no other anti-toxin gene found. Also some hits for FIC domain nucleotidyl transferase, but rquired HPFxxGNGR motif was not present. NKF for now.
Stop **77015**: Not sure if real gene. Strong coding potential present on GM host, but neither Aleemily or Cafasso call this gene.
Region upstream (left) of reverse gene stop @ **75958**: in Aleemily there is an orpham gene called here, but not in Cafasso. I reviewed GM self and host, and found no strong coding potential, so have left this region as a gap.

Please record yes/no for each of the questions below. If further explanation is needed, please add this item to the above box.

In the submitted DNA Master file (Yes/No):

Yes 1. Does the genome sequence in your submitted DNA Master file match the nucleotide fasta file posted on phagesDB (same number of bases, no N bases, etc.)?
Yes 2. Are all the genes 'Valid" when you click the Validation button?
Yes 3. Are the genes (and matching LocusTag numbers) sequential, starting with #1, counting by 1s.
Yes 4. Are the Locus Tags the "SEA_PHAGE NAME" format?
Yes 5. Has the documentation been recreated from the Feature Table to match the latest file version?
n/a 6. Have tRNAs followed the tRNA protocol, **COPYING** tRNA-AMINOACID type (DNA equivalent of the anti-codon) from Aragorn output - tRNA-Gln(ctg) - AND the ends been adjusted to match the Aragorn output?
Yes 7. Has the frameshift in the tail assembly chaperone been annotated correctly (if applicable)?
NO 8. Have you cleared your Draft_Blast data and have you re-Blasted the submitted DNA Master file?
I was unable to re-BLAST within DNA master and have yet to figure out what the issue is. Debbie said it would be fine for now to note this here and submit.
Yes 9. Has every gene been described and supported in your Supporting Data file?
Yes 10. Did you investigate 'gaps'?
Yes 11. Did you delete the genes that you meant to delete?

Now, make a profile of the file you plan to send. (And you can save this file for Review to Improve!)

Yes 1. Have any duplicate genes been deleted?
Yes 2. Has the Notes field been cleared (using the automated buttons)?
Yes 3. Do the gene numbers and locus tags match?
Yes 4. Are the correct Feature_Types correctly selected (most will be ORFs, but check that tRNAs and tmRNAs are correctly labeled)?
Yes 5. Do the function names in the Product field either match the official function list or say "Hypothetical Protein"?
Yes 6. Has the Function field been cleared (using the automated buttons)?

How are you documenting your gene calls in class? Choose any/all that apply:

X PECAAN output
DNA Master shorthand (previously used format)
Spreadsheet
Powerpoint
Word document (must be easily searchable)
Other: Describe.

What is the file type (sort) submitted for QC to document your gene calls? Choose only one.:

X PECAAN output
DNA Master shorthand (previously used format)
Spreadsheet
Powerpoint
Word document (must be easily searchable)
Other: Describe.